Multi-period conditional distribution functions for heteroscedastic models with applications to VaR.

Raymond Brummelhuis, Dominique Guegan

£2919

Multi-period Conditional Distribution Functions for Heteroscedastic Models with Applications to VaR

Raymond Brummelhuis [1]   **Dominique Guégan** [2 3]

Abstract

Let $(r_t)_t$ be a GARCH(1, 1) process with time-dependend variance σ_t^2. We study, for an arbitrary but *fixed* k, the large deviation asymptotics of r_{t+k} and of $r_{t+1} + \cdots + r_{t+k}$, given r_t, σ_t and their consequences for extreme quantile estimation. If $(r_t)_t$ models the time series of daily returns of some financial asset, our results are relevant for the estimation of multi-period Value at Risk, and show, in particular, that the heuristic "$\sqrt{k}$-rule" used in financial risk management is false in the context of these GARCH models.

Keywords: generalized autoregressive heteroskedastic process, conditional probability density functions, large deviation probabilities, asymptotics, Laplace integrals, quantile estimation, Value at Risk.

1 Introduction

The Value at Risk- or VaR associated with a position taken in the financial market can loosely be defined as the maximum expected loss within a chosen confidence and over a chosen time-frame. It's specification therefore requires the following input:

[1]University of Reims, UPRESA 6056 -. E-mail : raymond.brummelhuis@univ-reims.fr

[2]University of Reims, UPRESA 6056 -. E-mail : dominique.guegan@univ-reims.fr

[3]The authors gratefully acknowledge the support of the Fondation de la Banque de France for this research, through the project "Mesures de risques extrêmes en finance et choix du temps" of the "appel d'offres nr. 4"

- a time window $[t, t+k]$.
- a confidence level c, typically close to 1.
- a model for the behavior of (the financial assets making up) one's position over the chosen time frame.

With these data, VaR is simply the $(1-c)$-th lower quantile of the probability distribution of the Profit & Loss (P & L)-function over the period $[t, t+k]$; cf. RiskMetrics [22] and also Jorion [18], Dowd [11]. Time-frame and confidence level are simply parameters, which are at the user's discretion : choices for c of 95% or 99% and of time frames with k equal to 1 to 10 days are current, the one of $k = 10$ and $c = 0.95$ figuring in the Bank of International Settlements' 1996 capital adequacy requirements. We note here that time will be discrete in this paper, measured in days or in multiples of some other basic unit.

The choice of model is of course crucial. Here, "model" is to be understood in a wide sense: ranging from straightforward Historical Simulation or HS to more or less sophisticated parametric models. The disadvantages of HS are well-known: unreliable small quantile estimation due to lack of sufficient data for extreme events, the assumption that the future will be exactly like the past and too much weight given to distant (in time) events[4]. Parametric models of course come with their own dangers, the most obvious one of which is trying to make the data fit a into straightjacket unsuitable for them. For example, a lot of VaR methodologies are still based on (conditionally) normally distributed returns, although the deviation of empirical stock return data from the normal distribution, in particular as concerns the outliers, is by now well known and well documented: cf. Embrechts, Klüppelberg and Mikosch [12], Dowd [11], Frey and McNeil [15], Mikosch and Stărică [21]. Still, parametric models present the advantage of providing easily evaluated and more reliable extreme quantile-estimates (if the model's fit to the data is right, that is).

It is clear that VaR-estimates can vary hugely from one model to another, and that model risk is an important issue here. It is therefore important to thoroughly understand the mathematical implications of working with a certain stochastic

[4]The HS-method makes a hidden assumption that the P& L-process is stationary (which is of course one way of formalising that the "future is like the past") and non-parametrically estimates the *unconditional* probability distribution of the P&L process. Even if this process were in general stationary, conditional and unconditional probability distributions may differ hugely, and it is the conditional distributions which are important for day to day risk management, as has been stressed by Frey and McNeil [15].

model. Related to this, there also exists a different kind of model risk,which consists of naively applying conclusions valid for one model to other, mathematically speaking completely different ones, without any a priori or a posteriori justification. This would almost seem to be too obvious a point to make, but is exactly the kind of thing which can be observed in some of the VaR literature, and in particular in the RiskMetrics methodology for estimating VaR: cf. [22]. The RiskMetrics model is a particular example of the GARCH(1, 1)-models which are the main subject of this paper. It takes the form:

$$\begin{aligned} r_{t+1} &= \sigma_{t+1}\varepsilon_{t+1} \\ \sigma_{t+1}^2 &= \lambda\sigma_t^2 + (1-\lambda)r_t^2, \end{aligned} \tag{1}$$

where $r_t = \log(P_t/P_{t-1})$ is the logarithmic one-period return, P_t being the (monetary) value of one's position at time t, $\lambda \in (0,1)$ is a parameter to be estimated and where the ε_n are iid $N(0,1)$-distributed random variables. The motivation for (1) comes from the Exponential Moving Average Method for estimating daily volatilities, as explained in [11], [18] and [22]. In this model, the VaR over the period $[t, t+1]$ is easily computed to be

$$VaR_{1-c}(1) = -\sigma_{t+1} q_{1-c}^N P_t,$$

where q_α^N denotes the lower α-th quantile of the standard normal distribution[5]. We will often write $1-c = \alpha$. Note that it is the *conditional* VaR we are talking about. One next is interested in the k-period VaR. Following RiskMetrics, one easily computes that for any $\nu \geq 1$ the conditional expectation of $r_{t+\nu}^2$ given r_t and σ_t is simply σ_{t+1}^2 again, and that therefore the (conditional) variance of the k-period return is equal to:

$$\mathbf{E}\left((r_{t+1} + \cdots + r_{t+k})^2 | r_t, \sigma_t\right) = \mathbf{E}\left(r_{t+1}^2 + \cdots + r_{t+k}^2 | r_t, \sigma_t\right) = k\sigma_{t+1}^2,$$

the expectation being the one conditional to the values of r_t and σ_t at time t; note that we used in the first equality that the r_t's are uncorrelated. RiskMetrics then proposes to simply compute the k-period VaR over $[t, t+k]$ as

$$VaR_{1-c}(k) = \sqrt{k}VaR_{1-c}(1). \tag{2}$$

However, this only makes sense if the k-period returns are (close to) normally distributed[6] and, as one of the main results of our paper shows, this is far from

[5] For simplicity we make the usual approximation $e^r - 1 \simeq r$; this is of course not essential.

[6] In fact, (2) is the same as the k-period VaR in a simple random walk model for the r_t's, which should be enough to make one suspicious!

being the case, even for a k as small as 2. Therefore, baring accidental anumerical coincidences for specific c, one should expect the real VaR of the model to be very different from (2). Phenomena of this kind have been observed numerically, through Monte-Carlo simulation, by Frey and McNeil [15] (but for a different model, namely an AR(1)-GARCH(1, 1) with non-normal innovations ε_t estimated from empirical data).

A similar preoccupation with variances can be observed in much of the empirical financial and econometrical litterature. However, variance by itself is a poor indicator of risk if nothing is known about the underlying probability distribution, and primary attention should be given to the latter. It is one of the merit's of VaR, whatever it's adequacy as a risk management tool (cf. Artzner, Delbaen, Eber, and Heath [1]) that it does exactly that. Note, incidentally, that for a riskmeasure like the expected shortfall, which is defined by

$$e_\alpha(k) = \mathbf{E}(P_{t+k} - P_t | P_{t+k} - P_t \leq VaR_\alpha(k))$$

the dependence on the probability distribution might even be larger. Drawing conclusions from variances or standard deviations strictly speaking only makes sense when all relevant (cumulative) probability distributions belong to a one-parameter family $F(x/\sigma)$ (where in practice F might be allowed to vary slowly with σ). As we will see, this is definitely not the case for the different multi-period returns in a GARCH-model.

Autoregressive Conditionally Heteroskedastic or ARCH processes were introduced by Engle [13] and subsequently generalized by Bollerslev [2] to GARCH or Generalized ARCH processes, which are processes of the form (1), but with a σ_{t+1} which now more generally is given by

$$\sigma_{t+1} = (a_0 + \sum_{j=1}^{p} a_j r_{t-j}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2)^{1/2}.$$

The $(\varepsilon_t)_t$ are still supposed to be iid, but not necessarily normal. Such a process is called a GARCH(p, q). An ARCH(p) corresponds to the case of $q = 0$, or no dependence on past volatilities. More general processes allowing more general functional dependencies of σ_t on past values r_{t-j} of the process, and on past variances were introduced by for example Nelson [20], Guégan and Diebold [7], [8] and many others: we refer to Bollerslev, Engle and Nelson [3] or Guégan [16] for an overview. We will always assume that our ε_t have a probability density, which implies that the r_t and σ_t will have one also, at least conditionally.

GARCH processes are popular in empirical finance because of their capability to model phenomenae like volatility clustering and heavy tails (for the unconditional distribution, even if the ε_t's do not have one). One very often restricts oneself, as we will do here, to GARCH(1, 1) models, which combine a large descriptive force with a small number of parameters to be estimated. A lot of the theoretical work on GARCH-processes has concentrated on stationarity issues, in particular the existence and tail behavior of the stationary distribution: see for example Embrechts, Klüppelberg and Mikosch [12] or Mikosch and Stărică [21]. In this paper we will take a closer look at the distribution functions of the "k-days into the future" returns r_{t+k}, and of the cumulated returns $r_{t+1} \cdots + r_{t+k}$, for an arbitrary but *fixed* k, conditional to given values of r_t and σ_t at time t. Assuming we have standard normal "shocks" ε_t, we will determine the asymptotic behavior of the conditional probability densities of the former for large values of their arguments. For r_{t+k} these asymptotics will be shown to be of the type form

$$\mathbf{P}(r_{t+k} = x | r_t, \sigma_t) \simeq C_k |x|^{-(1-1/k)} e^{-c_k |x|^{2/k}},$$

with explicit constants c_k, C_k. For the cumulated returns we have only arrived at asymptotic upper and lower bounds of this form, with different constants in the two cases. Note that it is exactly this kind of asymptotics, for a fixed k and large $|x|$, which might be expected to be relevant for day-to-day risk management in a conditional probability setting. Also note the huge differences, in tail behavior, of these distributions for different k. In fact, it follows immediately that for $k \geq 2$ they fall into the class of subexponential distributions (cf. [12]).

It is of course to be expected that the probability distribution of r_{t+k} will differ very much from the normal one, for big k, since, under suitable hypothesis on the coefficients a_0, a_1 and b_1, the former will tend to the stationary distribution, which is known to be heavy-tailed (cf. [12], [21] and their references to earlier work, in particular Kesten [19]). What may be unexpected is that this deviation from normality already shows up that strongly for a k as small as $k = 2$. Incidentally, the above asymptotics and known tail estimates for a stationary GARCH(1, 1) imply that the $|x| \to \infty$ and $k \to \infty$ limits do not commute (which one had no right to expect to anyhow).

We stress that stationarity issues will not play a rôle in this paper, since we will only work at a fixed k. Observe, incidentally, that the RiskMetrics model (1) is neither second order nor strongly stationary (cf. [21]).

The paper is organized as follows: first, in section 2, we derive a general representation formula for the probability density function of r_{t+k}, valid for a general GARCH(1, 1) with rather arbitrary dependence of σ_{t+1} on r_t and σ_t. The classical GARCH(1, 1) of Bollerslev [2] and the EGARCH(1, 1) of Nelson [20] will provide illustrations. In section 3 we derive a similar formula for the cumulative returns $r_{t+1} + \cdots + r_{t+k}$. The remainder of the paper will then be concerned with the asymptotics of these distribution functions in the case of a classical GARCH(1, 1) with normal innovations[7]. In section 4 we first derive a technical result on asymptotics of Laplace transforms which will be needed to analyze these asymptotics. In section 5 we then obtain rather precise asymptotics for the probability density functions of the returns r_{t+k} and, in section 6, a somewhat more qualitative result for the cumulative returns. We end the paper with a section on applications to quantile- or VaR estimation.

The main results of this paper were announced in [5].

2 "k-step into the future" conditional distributions for GARCH(1, 1) processes

We consider a general $GARCH(1,1)$ process of the form:

$$\begin{cases} r_{t+1} = \sigma_{t+1}\varepsilon_{t+1} \\ \sigma_{t+1} = \varphi(r_t, \sigma_t) \end{cases} \tag{3}$$

We will typically have in mind the case of r_t modeling some security return: $r_t = \log(P_t/P_{t-1})$, with $(P_t)_{t\in\mathbf{N}}$ is some time series of asset prices.

The function $\varphi : \mathbf{R}^2 \to \mathbf{R}_{\geq 0}$ will be some measurable function, which will have to satisfy an additional condition to be specified below. As for the random shocks, or innovations, $(\varepsilon_t)_t$, we will make the usual hypothesis that they be iid, with mean 0 and variance 1. This could be weakened to independence only, and one could also let φ explicitly depend on t, but since such changes can easily be incorporated afterwards, we will first concentrate on the iid case. Independence of the ε_t will however be essential for us, and the results of this paper are not expected to go through without change for, for example, the important class of weak GARCH-processes introduced by Drost and Nijman (although these would have been very convenient for us when discussing multi-period returns in section 6 below, this class being closed under aggregation, unlike ordinary GARCH).

We will assume that the probability distribution functions of the random variables ε_t and $\varphi(u\varepsilon_t, u)$, for arbitrary but fixed $u > 0$, have a density. This condition can probably also be weakened in many places, but since most GARCH-models used in practice satisfy this requirement, we decided to limit ourselves to this case. We will use the notation $X \sim f$ to indicate that the random variable X has

[7]Non-normal innovations will be considered in another paper

probability density f and will sometimes use the abreviation "pdf" for probability density function. We therefore assume that there exist L^1-functions f and h_u, for each $u > 0$, such that

$$\begin{aligned} \varepsilon_t &\sim f & (4) \\ \varphi(u\varepsilon, u) &\sim h_u \text{ if } \varepsilon \sim f,\ u > 0. & (5) \end{aligned}$$

We note that (5) is in fact the additional hypotheses on φ alluded to above. As we will see below in some examples, the functions h_u can easily be computed in practice.

We will be interested, in this section and in much of this paper, in the conditional probabilities of r_{t+k}, given r_t and σ_t, where k is an arbitrary, but fixed, natural number. If we denote the underlying probability measure by $\mathbf{P}(\cdot)$, then we would like to understand the properties of the conditional pdf's

$$\begin{aligned} p_{t,k}(x;\rho,s) &:= \frac{d}{dx}\mathbf{P}\left(r_{t+k} \leq x | r_t = \rho_0, \sigma_t = s_0\right) \\ &=: \mathbf{P}\left(r_{t+k} \leq x | r_t = \rho_0, \sigma_t = s_0\right), \qquad (6) \end{aligned}$$

where the derivative is in the Radon-Nykodim sense, and where the somewhat formal "physicists'" notation introduced in the last line will be used quite often throughout the paper. Our process being, for the moment, time-homogeneous, one could of course take $t = 0$ here. However, in financial applications, when modeling for example daily returns, t will stand for "today", and ρ_0 and s_0 will be todays *observed* return and volatility, which will change in going from today to tomorrow, so to speak. So there is still some dynamic time dependence present, hidden in the ever changing initial conditions (ρ_0, s_0), and to retain this dynamical flavor, if only for psychological reasons, we prefer to keep the time t explicit in our notations.

We next define two integral operators F and H by

$$\begin{aligned} F(u)(x) &= \int_0^\infty \frac{1}{s} f\left(\frac{x}{s}\right) u(s)\ ds, & (7) \\ H(u)(s) &= \int_0^\infty h_{s'}(s) u(s') ds', & (8) \end{aligned}$$

acting on $L^1(\mathbf{R})$ and $L^1(0,\infty)$, respectively; f and $h_{s'}(\cdot)$ being pdf's, these operators are positivity-preserving, and an easy application of Fubini's theorem shows that they are well-defined and have norm 1: for example,

$$||F(u)||_1 \leq \int\int \frac{1}{s} f\left(\frac{x}{s}\right) |u(s)| ds dx$$

$$
\begin{aligned}
&= \int \left(\int \frac{1}{s} f\left(\frac{x}{s}\right) dx \right) |u(s)| ds \\
&= ||u||_1,
\end{aligned}
$$

with equality throughout if u is non-negative. We note that for, say, continuous f and h_u's, F and H can also be defined on the spaces of finite Radon measures on $\mathbf{R}$ and $\mathbf{R}_{>0}$, respectively, and that they are also positivity-preserving operators of norm 1 on these. We will find it occasionally notationally convenient to let them formally act on delta measures, as for example in the following theorem; this will however never be essential, and can always be avoided.

We now can state the main result of this section:

Theorem 2.1 *(Markovian representation formula for $p_{t,k}$.) Let $(r_t)_{t\in\mathbf{N}}$ be defined by (3). Under the above hypotheses on ε_t and φ, we then have that*

$$
\begin{aligned}
p_{t,k}(x; \rho_0, s_0) &= \mathbf{P}\left(r_{t+k} = x | r_t = \rho_0, \sigma_t = s_0\right) && (9) \\
&= F \circ H^{k-1}\left(\delta_{\varphi(\rho,s)}\right). && (10)
\end{aligned}
$$

Written out explicitly, if $k > 1$, then

$$
p_{t,k}(x : \rho, s) = \int_{(R_{\geq 0})^{k-1}} \frac{1}{s_k} f\left(\frac{x}{s_k}\right) h_{s_{k-1}}(s_k) h_{s_{k-2}}(s_{k-1}) \cdots h_{\varphi(\rho,s)}(s_2) ds_2 \cdots ds_k. \tag{11}
$$

Proof. If $k = 1$, the pdf of r_{t+1}, given $r_t = \rho_0$ and $\sigma_t = s_0$ is clearly equal to

$$
\frac{1}{\varphi(\rho_0, s_0)} f\left(\frac{x}{\varphi(\rho_0, s_0)}\right)
$$

which proves (9) for this case. If $k > 1$ then

$$
\begin{aligned}
&\mathbf{P}\left(r_{t+k} = x | r_t = \rho_0, \sigma_t = s_0\right) \\
&= \mathbf{P}\left(\sigma_{t+k}\varepsilon_{t+k} = x | r_t = \rho_0, \sigma_t = s_0\right) \\
&= \int_0^\infty \mathbf{P}\left(\sigma_{t+k}\varepsilon_{t+k} = x | \sigma_{t+k} = s_k, r_t = \rho, \sigma_t = s\right) \cdot \mathbf{P}\left(\sigma_{t+k} = s_t | r_t = \rho_0, \sigma_t = s_0\right) ds_k \\
&= \int_0^\infty \frac{1}{s_k} f\left(\frac{x}{s_k}\right) \cdot \mathbf{P}\left(\sigma_{t+k} = s_k | r_k = \rho_0, \sigma_t = s_0\right) ds_k, && (12)
\end{aligned}
$$

since ε_{t+k} is independent of $\sigma_{t+k} = \varphi(r_{t+k-1}, \sigma_{t+k-1})$. Next, $r_{t+k-1} = \sigma_{t+k-1}\varepsilon_{t+k-1}$, the two factors on the right hand side being independent again. Hence

$$
\mathbf{P}\left(\sigma_{t+k} = s_k | r_k = \rho_0, \sigma_t = s_0\right)
$$

$$
\begin{aligned}
= & \int_0^\infty \mathbf{P}\left(\varphi(s_{k-1}\varepsilon_{t+k-1}, s_{k-1}) = s_k | \sigma_{t+k-1} = s_{k-1}, r_t = \rho_0, \sigma_t = s_0\right) \\
\cdot & \mathbf{P}\left(\sigma_{t+k-1} = s_{t-1} | r_t = \rho_0, \sigma_t = s_0\right) ds_{k-1} \\
= & \int_0^\infty \mathbf{P}\left(\varphi(s_{k-1}\varepsilon_{t+k-1}, s_{k-1}) = s_k\right) \cdot \mathbf{P}\left(\sigma_{t+k-1} = s_{k-1} | r_t = \rho_0, \sigma_t = s_0\right) ds_{k-1} \\
= & \int_0^\infty h_{s_{k-1}}(s_k) \cdot \mathbf{P}\left(\sigma_{t+k-1} = s_{k-1} | r_t = \rho_0, \sigma_t = s_0\right) ds_{k-1}.
\end{aligned}
$$

Substituting this expression in (12) and repeating the argument for $P\left(\sigma_{t+k-1} = s_{k-1} | r_t = \rho_0, \sigma_t = s_0\right)$ we obtain, after k steps, formula (11), which proves the theorem. QED

Remarks 2.2 (i) We should note that for an ARCH(1), by which we will mean a process of the form (3) with a function φ which does not depend on σ, one can immediately write down a simpler formula, which is a direct consequence of the Markov property of the process:

$$
p_{t,k}(x; \rho_0) = F^k(\delta_{\rho_0}), \tag{13}
$$

where the integral operator F is defined by:

$$
F(u)(x) = \int_{\mathbf{R}} \frac{1}{\varphi(y)} f\left(\frac{x}{\varphi(y)}\right) dy;
$$

see also remark at beginning of section 3.

(ii) We can make the model (3) a little more realistic, as a stock return model, by adding a mean which is a function of the previous day's volatility:

$$
\begin{aligned}
r_{t+1} &= \mu_{t+1} + \sigma_{t+1}\varepsilon_{t+1} \\
\mu_{t+1} &= \psi(\sigma_t)
\end{aligned}
$$

and σ_{t+1} given by $\varphi(r_t, \sigma_t)$, as before. Such a model, with ψ an affine function of σ: $\psi(\sigma) = r + \beta\sigma$, was introduced by Engle, Lilian and Robbins (1987), and is called an ARCH-M model. It can be understood as inspired by the Capital Asset Pricing Model or CAPM from Finance, in which an investor's expected return is the sum of some riskless returm, r, and a term which is proportional to the stock's risk, as measured by it's standard deviation or volatility. The ARCH-M model is easily incorporated in theorem 2.1, simply by replacing the kernel of F in (7) by

$$
\frac{1}{s} f\left(\frac{x - \psi(s)}{s}\right).
$$

Examples 2.3 To illustrate the use of theorem 2.1 we look at some examples.

(i) **Classical GARCH(1, 1):** We take

$$\varphi(r,\sigma) = \left(a_0 + a_1 r^2 + b_1\sigma^2\right)^{1/2}, \tag{14}$$

and ε_t iid, $\varepsilon_t \sim f$. We leave f unspecified, apart form requiring that it will have mean 0 and variance 1. We can easily compute the kernel $h_u(s)$ in terms of f: if $\varepsilon \sim f$, then the pdf of $\varphi(u\varepsilon, u)$ is:

$$\frac{d}{ds}\mathbf{P}\left(\left(a_0 + a_1u^2\varepsilon^2 + b_1u^2\right)^{1/2} < s\right),$$

which will be equal to 0 if $s \le \sqrt{a_0 + b_1u^2}$, and equals

$$\frac{d}{ds}\left(\int_{-(s^2-a_0-b_1u^2/a_1u^2)^{1/2}}^{(s^2-a_0-b_1u^2/a_1u^2)^{1/2}} f(y)\,dy\right) =$$

$$\frac{1}{2}s\left(a_1u^2(s^2-a_0-b_1u^2)\right)^{-1/2}\sum_{\pm} f\left(\pm\left(\frac{s^2-a_0-b_1u^2}{a_1u^2}\right)^{1/2}\right),$$

if $s > \sqrt{a_0 + b_1u^2}$. If f is symmetric, which in applications is often the case, this simplifies to:

$$2s\left(a_1u^2(s^2-a_0-b_1u^2)\right)^{-1/2} f\left(\left(\frac{s^2-a_0-b_1u^2}{a_1u^2}\right)^{1/2}\right)\chi_{\{s>\sqrt{a_0+b_1u^2}\}}, \tag{15}$$

χ_A being the indicator function of a set A. Popular choices for f are the standard Gaussian distribution

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2},$$

and, more recently also heavy tailed distributions like a Student t-distribution with a low number of degrees of freedom, to account for the empirically observed heavy tails of many financial asset returns. Frey and McNeil [15] use Extreme Value Theory to estimate the tails of f and a non-parametric distribution for it's center. As mentioned in the introduction, in this paper we will concentrate on standard normal innovations ε_t.

We note that theorem 2.1, together with formula (15), can be used for fast numerical computation of the probability densities (6), as an alternative to Monte-Carlo: we refer to [6] for examples.

(ii) **EGARCH(1, 1):** In Nelson's Exponential GARCH or EGARCH(p, q) model, σ_{t+1} is given by:

$$\left(1+\sum_{j=1}^{p}\beta_j L^j\right)\log\sigma_{t+1}^2=\omega+\left(1+\sum_{i=1}^{q}\alpha_i L^i\right)g(\varepsilon_t),$$

where

$$g(\varepsilon)=\gamma\left(|\varepsilon|-\mathbf{E}(|\varepsilon|)\right)+\theta\varepsilon,$$

and where L denotes the usual lag or back-shift operator: see Nelson [20] or Bollerslev, Engle and Nelson [3], Guégan [16]. If $q=0$ and $p=1$, this is again a special case of (3), with

$$\varphi(r,\sigma)=\sigma^{-\beta}e^{(\omega+g(r/\sigma))/2}.$$

It is again straightforward, though a little tedious, to compute the kernel $h_{s'}(s)$, and we only state the result. Let $C:=\omega-\mathbf{E}(|\varepsilon|)=\omega-\int|y|f(y)dy$. Then, assuming that for example $\theta>\gamma>0$,

$$\begin{aligned} h_u(s) &= \frac{2}{(\theta+\gamma)x}f\left(\frac{2(\log s+\beta\log u)-C}{\theta+\gamma}\right)\chi_{[e^{C/2}u^{-\beta},\infty)}(s)\\ &+ \frac{2}{(\theta-\gamma)x}f\left(\frac{2(\log s+\beta\log s')-C}{\theta-\gamma}\right)\chi_{[0,e^{C/2}u^{-\beta})}(s)\end{aligned}$$

There exist numerous other GARCH models in the litterature (cf. for example [3], [16], [17], Dingh and Granger [9], Diebolt and Guégan [7], [8]: this list is far from being exhaustive!), for which similar computations can be carried out.

(iii) **Moments:** As a final illustration of formula (9), we show how it can be used to compute the moments, and, more generally, the conditional expectation of any $g(r_{t+k})$ for sufficiently "nice" g. In fact,

$$\begin{aligned} &\mathbf{E}\left(g(r_{t+k})|r_t=\rho,\sigma_t=s\right)\\ =&\int_{\mathbf{R}}\int_{\mathbf{R}^k_{\geq 0}}\frac{1}{s_k}f\left(\frac{x}{s_k}\right)h(x)\left(\Pi_{\nu=2}^{k}h_{s_{\nu-1}}(s_\nu)\right)\delta_{\varphi(\rho,s)}(s_1)ds_1\cdots ds_k dx\\ =&\int_{\mathbf{R}^k_{\geq 0}}\mathbf{E}\left(g(s_k\varepsilon)\right)\left(\Pi_{\nu=2}^{k}h_{s_{\nu-1}}(s_\nu)\right)\delta_{\varphi(\rho,s)}(s_1)ds_1\cdots ds_k,\end{aligned}$$

ε being a random variable with pdf f. Taking $g(x)=x^n$ we therefore find that the n-th moment of r_{t+k} equals:

$$\mathbf{E}(r_{t+k}^n|r_t=\rho,\sigma_t=s)=\mu_{n,f}\int_{\mathbf{R}^k_{\geq 0}}s_k^n\left(\Pi_{\nu=2}^{k}h_{s_{\nu-1}}(s_\nu)\right)ds_2\cdots ds_k, \qquad (16)$$

where $\mu_{n,f} = \int x^n f(x)dx$ denotes the n-th moment of f and where we put $s_1 := \varphi(\rho_0, s_0)$. Note that this formula only involves the kernels $h_{s'}(s)$. As is well-known, for a classical GARCH(1, 1), with φ given by (14), the even moments can be calculated recursively, using the special form of the φ; note, however, that this will not work of a general GARCH(1, 1) with arbitrary φ and neither will give information on the odd moments, if f is not symmetric.

We end this section by indicating how to adapt formula (11) when both φ and the pdf of ε_t depend explicitly on t: $\varepsilon_t \sim f_t$, say. In that case we simply define the kernels $h_u^t(s)$ by

$$\varphi(u\varepsilon_t, u, t) \sim h_u^t(\cdot) \quad \text{if} \quad \varepsilon_t \simeq f_t$$

(assuming, as before, that the left hand side has a pdf) and replace $h_{s_{j-1}}(s_j)$ in (11) by $h_{s_{j-1}}^{t+j-1}(s_j)$ and $s_k f(x/s_k)$ by $s_k f_{t+k}(x/s_k)$.

Letting φ and the pdf of ε_t depend on time might have some relevance for practical modeling purposes, but will of course pose numerous identification and estimation problems. We are not aware of empirical work where this has been done.

3 Multiple period returns

Estimating Profit & Loss over a multiple period time window $[t, t+k]$ involves looking at the k-period return

$$r_{t+k,t} = \log(P_{t+k}/P_t) = r_{t+1} + \cdots + r_{t+k}, \tag{17}$$

rather than just simply r_{t+k}. It is known that GARCH(1, 1) is not closed under agregation, cf. Drost and Nijman [10], so that we cannot expect the process $(r_{t,t+k})_t$ to be GARCH any more. The results of the previous section are therefore not immediately applicable, and we will proceed somewhat differently, by loking at the two-component Markov process $Z_t = (r_t, \sigma_t)$ under (3). We write $z_0 = (\rho_0, s_0)$ and we let P_{z_0} be the probability conditional to $Z_t = z_0$. Then the joint pdf of $(Z_{t+1}, \cdots, Z_{t+k})$ with respect to the probability $\mathbf{P}_{z_0}$ can be evaluated as:

$$\begin{aligned}
&\mathbf{P}_{z_0}\left((Z_{t+1}, \cdots, Z_{t+k}) = (z_1, \cdots, z_k)\right) = \\
&\Pi_{j=1}^k \mathbf{P}_{z_0}\left(Z_{t+j} = z_j | (Z_{t+1}, \cdots, Z_{t+j-1}) = (z_1, \cdots, z_{j-1})\right) = \\
&\Pi_{j=1}^k \mathbf{P}_{z_0}\left(Z_{t+j} = z_j | Z_{t+j-1} = z_{j-1}\right),
\end{aligned}$$

by the Markov property. It follows that this joint pdf equals:

$$\mathbf{P}\left((r_{t+j}, s_{t+j}) = (x_j, s_j),\ 1 \le j \le k | (r_t, s_t) = (\rho_0, s_0)\right) =$$
$$\Pi_{j=1}^{k} \frac{1}{s_j} f\left(\frac{x_j}{s_j}\right) \delta(s_j - \varphi(x_{j-1}, s_{j-1})), \tag{18}$$

$\delta(s - v)$ being the Dirac delta measure and $s_0 = \varphi(\rho_0, s_0)$; the occurence of these Dirac measures is natural, since (3) is deterministic in the second component. Note that this would be different for a stochastic volatility type of model, which can in principle be treated in the same way.

The conditional pdf of $r_{t+k,t} = x$ now is found by integrating (18) against $\delta(x-(x_1+\cdots+x_k))$. We can evaluate the $s_1, \cdots, s_k$ - integrals involving the delta -functions and then obtain the following result:

Theorem 3.1 *Inductively define functions* $\hat{s_j} = \hat{s_j}(x_1, \cdots, x_{j-1})$ *by:*

$$\begin{aligned} \hat{s}_1 &= \varphi(\rho_0, s_0) \\ \hat{s}_j &= \varphi(x_{j-1}, \hat{s}_j) \end{aligned}$$

Then

$$\mathbf{P}(r_{t+k,t} = x | r_t = \rho_0, \sigma_t = s_0) =$$
$$\int \cdots \int \frac{1}{\hat{s}_k} f\left(\frac{x - (x_1 + \cdots + x_{k-1})}{\hat{s}_k}\right) \Pi_{j=1}^{k-1} \frac{1}{\hat{s}_j} f\left(\frac{x_j}{\hat{s}_j}\right) dx_1 \cdots dx_{k-1}. \tag{19}$$

Remark 3.2 It is possible to rederive theorem 2.1 along these lines, by integrating (18) over everything except x_k and making suitable changes of variables.

Note that formula (19) is not a simple operator product anymore. This complicates both the analysis of it's asymptotics for large $|x|$ in section 6 below, as well as it's numerical implementation, since multi-dimensional integrals are in general difficult to handle numerically. It would be interesting to see to what extend modern (non-Monte Carlo) methods of numerical integration like those in Sloan and Joe [23] can handle integrals of the form (17) up to the $k = 10$ which is relevant for risk management practice and for the different probability densities f of interest (Gaussian, Student, Generalized Pareto). Evaluating (17) by simple Monte Carlo would not be very efficient, and would also unnecessarily complicate matters, since one then might just as well directly simulate the process $(r_n)_n$ and read of the desired quantile of $r_{t+k,t}$ from the empirical distribution.

4 Asymptotics of Laplace Transforms

In this section we prove the following technical lemma on asymptotics of Laplace integrals which will be the basis of the asymptotic results proved in section 5 and 6.

Lemma 4.1 Let $\alpha > 0$, $s > 0, c > 0$ and $\beta \in R$. Then the following asymptotic development holds as $s \to \infty$:

$$\int_0^\infty x^{-\beta} e^{-cx^{-\alpha}} e^{-sx} dx \simeq \tag{20}$$
$$\left(\frac{s}{\alpha}\right)^{\frac{\beta-(\alpha/2)-1}{\alpha+1}} e^{-(\alpha+1)c^{1/(\alpha+1)}(\frac{s}{\alpha})^{\alpha/\alpha+1}} \sum_{j=0}^\infty C_j s^{-j\alpha/(\alpha+1)},$$

with $C_0 = \sqrt{2\pi/\alpha(\alpha+1)} c^{(1-2\beta)/2(\alpha+1)}$.

Remarks 4.2 (i) The meaning of the sign $\simeq$ is the usual one: if we cut the sum after $N-1$ terms, then there exist for each $R > 0$ a constant $C_{N,R} = C_{N.R}(c, \alpha, \beta)$ such that the error we make can, for $s > R$, be estimated by:

$$C_{N,R} s^{\frac{\beta-(N+1/2)\alpha-1}{\alpha+1}} e^{-(\alpha+1)c^{1/(\alpha+1)}(\frac{s}{\alpha})^{\alpha/\alpha+1}} \tag{21}$$

We will in the following, for simplicity, only record the main term of the various asymptotic series we will encounter in the various proofs below, and write (20) as:

$$\int_0^\infty x^{-\beta} e^{-sx-x^{-\alpha}} dx \simeq \tag{22}$$
$$e^{-(\alpha+1)(\frac{s}{\alpha})^{\alpha/\alpha+1)}} \left(\sqrt{\frac{2\pi}{\alpha(\alpha+1}} \left(\frac{s}{\alpha}\right)^{\frac{\beta-(\alpha/2)-1}{\alpha+1}} + O\left(s^{\frac{\beta-(3\alpha/2)-1}{\alpha+1}} \right) \right),$$

often even leaving out the O-term altogether. This should cause no confusion.

(ii)A closely related result is (the abelian part of) de Bruijn's Tauberian theorem, theorem 4.12.9 in Bingham, Goldey and Teugels [4], in which for example the function $x^{-\alpha}$ in the exponent in (20) can be, much more generally, replaced by any function having such an asymptotic behavior (in the sense of regular variation) as $x \to 0$, but in which the statement is weakened to asymptotic equivalence of the logarithms. We would like to thank Paul Embrechts for calling our attention to this reference.

Proof of lemma 4.1: It suffices to prove (20) for $c = 1$, by a simple scaling argument. Note that we cannot directly apply Watson's lemma, since all the derivatives of $x^{-\beta}\exp(-sx - x^{-\alpha})$ are 0 in 0. We will first split the integral in two, as follows:

$$\begin{aligned} \int_0^\infty x^{-\beta} e^{-sx-x^{-\alpha}} dx &= \int_0^{s^{-1/(1+\alpha)}} + \int_{s^{-1/(1+\alpha)}}^\infty \\ &= I + II, \end{aligned} \tag{23}$$

noting that $sx = x^{-\alpha}$ precisely when $x = s^{-1/(\alpha+1)}$, and then analyze the two parts separately, using Laplace's method (or complex stationary phase). We start with the second integral, II. Making the change of variables $x = s^{-1/(1+\alpha)}y$, we get

$$II = s^{\frac{\beta-1}{\alpha+1}} \int_1^\infty e^{-s^{\alpha/(\alpha+1)}(y+y^{-\alpha})} y^{-\beta} dy,$$

which, apart from the factor in front, is a classical Laplace integral of the form

$$\int_1^\infty e^{-\lambda\varphi(y)} a(y) dy.$$

The main contribution to the asymptotics will come from the absolute minimum of the phase function $\varphi(y) = y + y^{-\alpha}$ on $[1, \infty)$ and/or from the boundary point $y = 1$. It is easily seen that $\varphi(y)$ has an absolute minimum on $R_{>0}$ at $y = y_c = \alpha^{1/(\alpha+1)}$. We distinguish three cases:

(i) $\alpha > 1$. In this case, $y_c \in (1, \infty)$ and we get a contribution

$$e^{-\lambda\varphi(y_c)} \left(\left(\frac{2\pi}{\lambda} \right)^{1/2} \frac{a(y_c)}{\varphi''(y_c)^{1/2}} + O\left(\lambda^{-3/2} \right) \right)$$

where the O-term stands for a complete asymptotic series in powers $\lambda^{-(1/2)-j}$. Computing $\varphi(y_c) = \alpha^{1/(\alpha+1)} + \alpha^{-\alpha/(\alpha+1)} = (\alpha+1)\alpha^{-\alpha/(\alpha+1)}$ and $\varphi''(y_c) = \alpha(\alpha+1)/\alpha^{(\alpha+2)/(\alpha+1)}$ and remembering the factor in front and the fact that $\lambda = s^{\alpha/(\alpha+1)}$, we find the following contribution to II:

$$\exp\left((\alpha+1) \left(\frac{s}{\alpha} \right)^{\alpha/(\alpha+1} \right) \left(\sqrt{2\pi/\alpha(\alpha+1)} \left(\frac{s}{\alpha} \right)^{\frac{\beta-(\alpha/2)-1}{\alpha+1}} + \cdots \right), \tag{24}$$

the dots indicating lower order terms. We have to compare this with the contribution from the boundary point $y_c = 1$, which is

$$e^{-2s^{\alpha/(\alpha+1}} \left((1-\alpha)^{-1} s^{\frac{\beta-\alpha-1}{\alpha+1}} + \cdots \right). \tag{25}$$

However, these will all be dominated by (24), as follows from the following elementary observation:

For all $\alpha > 0$:

$$(\alpha+1)\alpha^{-\alpha/(\alpha+1)} \leq 2 \tag{26}$$

with equality iff $\alpha = 1$.

To prove (26) we have to show that

$$\log(\alpha+1) - \left(\frac{\alpha}{\alpha+1}\right)\log\alpha \le \log 2$$

for $\alpha > 0$. Now the derivative of the left hand side equals $-\frac{\log\alpha}{(\alpha+1)^2}$ which is 0 iff $\alpha = 1$ and which is > 0 (< 0) if $\alpha < 1$ ($\alpha > 1$). Hence the right hand side has an absolute maximum in $\alpha = 1$, which equals $\log 2$. QED

We continue with the proof of lemma 4.1. We consider the two remaining cases for II:

(ii) $\alpha = 1$: The minimum y_c coincides with the boundary point, and we obtain 1/2 times (24).

(iii) $\alpha < 1$: In this case $y_c < 1$ and the asymptotics of II will be given by (25), since only $y = 1$ will contribute.

We will now repeat the analysis for the first integral in (23), I. We make the substitutions $x = s^{-1/(\alpha+1)}u^{-1}$ and find that

$$I = s^{(\beta-1)/(\alpha+1)} \int_1^\infty e^{-s^{\alpha/(\alpha+1)}(u^\alpha+u^{-1})} u^{\beta-2} du.$$

Now the phase function $\varphi(u) = u^\alpha + u^{-1}$ will have an absolute minimum in $u = u_c = \alpha^{-1/(\alpha+1)}$ and we compute, as before, that $\varphi(u_c) = (\alpha+1)\alpha^{-\alpha/(\alpha+1)}$ and that

$$\begin{aligned} \varphi''(u_c) &= \alpha(\alpha-1)\alpha^{-\left(\frac{\alpha-2}{\alpha+1}\right)} + 2\alpha^{\frac{3}{\alpha+1}} \\ &= \alpha^{2/(\alpha+1)}\left(\alpha(\alpha+1)\right)\alpha^{-\alpha/(\alpha+1)}. \end{aligned}$$

We now consider the same three cases as for II:

(i') $\alpha > 1$: Since $u_c < 1$, the only contribution to the asymptotics will come from the boundary point $u = 1$, which will give (25).

(ii') $\alpha = 1$: $u_c = 1$ and we get 1/2 times (24), as before.

(iii') $\alpha < 1$. Now $u_c > 1$ will give a contribution to the asymptotics of I, which turns out to be the same as (24) (but with $\alpha < 1$, of course). By lemma (26) this contribution will win again from that coming from the boundary point.

It now suffices to add up the asymptotics of I and II and observe that, once more, by (26), in cases $(i)+(i')$ and $(iii)+(iii')$ the contribution of the interior minimum will dominate that of the boundary point. QED

5 Precise large deviation asymptotics for $r_{t+k}|r_t$

For the remainder of this paper $(r_t)_t$ will be a classical GARCH(1, 1), with

$$\varphi(r,\sigma) = (a_0 + a_1 r^2 + b_1\sigma^2)^{1/2} \tag{27}$$

and normally distributed iid ε_t. Fix an $(\rho_0, s_0) \in \mathbf{R} \times \mathbf{R}_{>0}$. Our aim is to analyze the asymptotic behavior,as $|x| \to \infty$, of the conditional pdf's

$$p_k(x) := p_k(x; t, \rho_0, s_0) = \mathbf{P}(r_{t+k} = x | r_t = \rho_0, \sigma_t = s_0).$$

We stress that we will study this asymptotics for arbitrary but *fixed* k. Our strategy is to deduce this asymptotics inductively from theorem 2.1, via a chain of lemma's which determine how the asymptotic behavior of a function v is affected by application of the operators F and H introduced in section 2. Recall formula (15) for the kernel of H for a classical GARCH(1,1), where we take for f the standard normal density: $f(x) = (2\pi)^{-1/2}e^{-x^2/2}$.

Lemma 5.1 *Suppose that $v(s) \simeq Cs^\beta e^{-cs^\alpha}$ for $0 < s \to \infty$, where $\beta \in R$, $c > 0$ and $\alpha > 0$, suppose that φ is given by (27). Then:*

$$Hv(s) \simeq C' s^{(2\beta-\alpha)/(\alpha+2)} e^{-c' s^{2\alpha/(\alpha+2)}}, \quad s \to \infty,$$

where

$$c' = \frac{1}{2}(\alpha+2)(\alpha a_1)^{-\frac{\alpha}{\alpha+2}} c^{\frac{2}{\alpha+2}}$$

and

$$C' = \frac{2Ce^{\frac{b_1}{2a_1}}}{\sqrt{\alpha+2}}(c\alpha a_1)^{-\frac{\beta+1}{\alpha+2}}.$$

Proof. We first treat the case of an $ARCH(1)$: $b_1 = 0$, which is computationally somewhat easier. In this case

$$Hv(s) = \gamma(s)\int_0^\infty \frac{1}{t} e^{-(s^2-a_0)/2a_1t^2} v(t)dt,$$

where

$$\gamma(s) = \frac{1}{\sqrt{2\pi}} \frac{2s}{\sqrt{a_1(s^2-a_0)}},$$

for $s^2 > a_0$, while $Hv(s) = 0$ for $s^2 \le a_0$. Making the change of variables $z = 1/t^2$, we obtain for $s^2 > a_0$, putting $\tilde{s} = \frac{s^2 - a_0}{2a_1}$:

$$Hv(s) = \frac{1}{2}\gamma(s)\int_0^\infty e^{-\tilde{s}z}\frac{1}{z}v\left(\frac{1}{\sqrt{z}}\right)dz.$$

The integral on the right hand side is the Laplace transform of $z^{-1}v(z^{-1/2})$ evaluated in $(s^2 - a_0)/2a_1$, whose large s-behavior is completely determined by the small z-behavior of

$$\frac{1}{z}v\left(\frac{1}{\sqrt{z}}\right) \simeq Cz^{-(\beta/2)-1}e^{-cz^{-\alpha/2}}, \quad z \to 0,$$

where we used the hypothesis on v. Part (i) of the lemma now follows from lemma 4.1 and straightforward calculations. We use here that $\exp(-c(s^2 - a_0)^{\alpha/\alpha+2}) \simeq \exp(-cs^{2\alpha/\alpha+2})(1 + \sum_\nu c_\nu s^{-2\nu})$ as $s \to \infty$, since $\alpha/\alpha + 2 \le 1$; note that this would in fact be false if this exponent were bigger than 1.

The argument for a $GARCH(1,1)$ ($b_1 \neq 0$) is slightly more involved. First, in that case,

$$Hv(s) = \frac{2se^{b_1/2a_1}}{\sqrt{2\pi a_1}}\int_0^{\sqrt{(s^2-a_0)/b_1}}\frac{1}{\sqrt{t^2(s^2 - a_0 - b_1t^2)}}e^{-(s^2-a_0)/2a_1t^2}v(t)dt,$$

if $s^2 > a_0$ and $Hv(s) = 0$ otherwise. Note that the integral no longer extends over the whole of the positive reals, as it did for an $ARCH(1)$. Making the same change of variables $z = 1/t^2$ as before, and putting $\gamma_1(s) = 2se^{b_1/2a_1}/\sqrt{2\pi a_1}$, we obtain that

$$\begin{aligned} Hv(s) &= \frac{1}{2}\gamma_1(s)\int_{b_1/(s^2-a_0)}^\infty \sqrt{\frac{z}{s^2 - a_0 - b_1z^{-1}}}e^{-(\frac{s^2-a_0}{2a_1})z}\frac{1}{z}v(\frac{1}{\sqrt{z}})dz \\ &= \frac{1}{2}\gamma_1(s)\frac{1}{\sqrt{b_1}}\int_{\widetilde{s}^{-1}}^\infty \sqrt{\frac{z}{\widetilde{s}z - 1}}e^{-(\frac{b_1}{2a_1})\widetilde{s}z}\frac{1}{z}v\left(\frac{1}{\sqrt{z}}\right)dz, \end{aligned}$$

where we have put $\widetilde{s} = (s^2 - a_0)/b_1$, for notational convenience. One easily sees that the main contribution to the $\widetilde{s} \to \infty$-behavior of this integral will again come from the asymptotics of $z^{-1}v(z^{-1/2})$ as $z \to 0$, which, as before, is given by $Cz^{-(\beta/2)-1}\exp(-cz^{-\alpha/2})$. We therefore have reduced the problem to the asymptotics of the following integral:

$$\int_{\widetilde{s}^{-1}}^\infty \sqrt{\frac{z}{\widetilde{s}z - 1}}e^{-b_1\widetilde{s}z/2a_1}z^{-(\beta/2-1)}e^{-cz^{-\alpha/2}}dz \tag{28}$$

$$
\begin{aligned}
&= \widetilde{s}^{(\beta-1)/2} \int_1^\infty \sqrt{\frac{w}{w-1}} e^{-b_1 w/2a_1} w^{-(\beta/2)-1} e^{-c\widetilde{s}^{\alpha/2} w^{-\alpha/2}} dw \\
&= \frac{2}{\alpha} \widetilde{s}^{(\beta-1)/2} \int_0^1 \left(\frac{1}{1-y^{2/\alpha}} \right)^{1/2} e^{-b_1 y^{-2/\alpha}/2a_1} y^{(\beta/\alpha)-1} e^{-c\widetilde{s}^{\alpha/2} y} dy,
\end{aligned}
$$

where we have made the changes of variables $w = \widetilde{s}z$ and $y = w^{-\alpha/2}$. The final integral is again a Laplace transform, whose main order asymptotic behavior is the same as that of

$$
\int_0^1 y^{(\beta/\alpha)-1} e^{-b_1 y^{-2/\alpha}/2a_1} e^{-c\widetilde{s}^{\alpha/2} y} dy. \tag{29}
$$

We can replace the interval of integration by $[0, \infty)$, thereby making an error of the form $O\left(s^{power} \exp(-(\text{const})s^{\alpha}\right)$ which will be of lower order, since $2\alpha/(\alpha+2) < \alpha$ for $\alpha > 0$. The resulting integral is a Laplace transform of the kind studied in lemma 4.1, with c equal to $b_1/2a_1$, β replaced by $-(\beta/\alpha)+1$, α by $2/\alpha$ and s by $c\widetilde{s}^{\alpha/2}$. After some calculations we find the asymptotics. QED

Lemma 5.2 *Suppose that $v(s) \simeq Cs^{\beta}e^{-cs^{\alpha}}$ for $0 < s \to \infty$, where $\beta \in R$, $c > 0$ and $\alpha > 0$. Then:*

$$
Fv(x) \simeq C'|x|^{(2\beta-\alpha)/(\alpha+2)} e^{-c'x^{2\alpha/(\alpha+2)}}, \quad x \to \infty,
$$

where

$$
c' = \frac{1}{2}(\alpha+2)c^{\frac{2}{\alpha+2}}(\alpha)^{-\frac{\alpha}{\alpha+2}},
$$

and

$$
C' = \frac{2C}{\sqrt{\alpha+2}}(c\alpha)^{-\frac{\beta+1}{\alpha+2}}.
$$

Proof. By the definition of F, we have that

$$
\begin{aligned}
Fv(x) &= \frac{2}{\sqrt{2\pi}} \int_0^\infty \frac{e^{-\frac{x^2}{2s^2}}}{s} v(s) ds \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty \sqrt{u} e^{-\frac{ux^2}{2}} v(\frac{1}{\sqrt{u}}) \frac{1}{u^{3/2}} du \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{ux^2}{2}} \frac{1}{u} v(\frac{1}{\sqrt{u}}) du,
\end{aligned}
$$

making the change of variables $u = 1/s^2$. The integral on the right hand side is the Laplace transform of $u^{-1}v(u^{-1/2})$ evaluated in $\frac{x^2}{2}$, whose large u-behavior is completely determined by the small u-behavior of

$$\frac{1}{u}v\left(\frac{1}{\sqrt{u}}\right) \simeq Cu^{-(\beta/2)-1}e^{-cu^{-\alpha/2}}, \quad u \to 0,$$

by the hypothesis on v. Using lemma 4.1 again, the asymptotics of $Fv(x)$ follow by straightforward calculations. QED

We next derive the asymptotic behavior of $H^k(\delta_{\varphi(\rho_0,s_0)})$:

Lemma 5.3 *Let $\varphi_0 := \varphi(\rho_0, s_0) = (a_0 + a_1\rho_0^2 + b_1 s_0^2)^{1/2}$. Then for any $k \geq 1$, if $s \to \infty$,*

$$H^k\left(\delta_{\varphi(\rho_0,s_0)}\right)(s) \simeq C_k s^{-(1-1/k)} e^{-\frac{k}{2a_1\varphi_0^{2/k}}s^{2/k}}, \tag{30}$$

where

$$C_k = \frac{e^{\frac{kb_1}{2a_1}}}{\sqrt{2\pi a_1}}\sqrt{\frac{2^{k-1}}{k}\varphi_0^{-1/k}}.$$

Proof. Let us write $q_k(s) = H^k(\delta_{\varphi_0})$ and $\beta_k = -(1-1/k)$, , $c_k = k/2a_1\varphi_0^{2/k}$, $\alpha_k = 2/k$. Finally, let C_k be as in the statement of this lemma. Then we will show by induction that $q_k \simeq C_k s^{\beta_k}\exp(-c_k s^{\alpha_k})$, for $s \to \infty$. First of all, for $k = 1$:

$$\begin{aligned} q_1(s) &= H(\delta_{\varphi(\rho_0,s_0)})(s), \\ &= \frac{1}{\sqrt{2\pi}}\frac{2s}{\sqrt{a_1\varphi_0^2(s^2 - a_0 - b_1\varphi_0^2)}}e^{-\frac{1}{2}\frac{s^2-a_0-b_1\varphi_0^2}{a_1\varphi_0^2}} \\ &\simeq \frac{1}{\sqrt{2\pi a_1}}\frac{1}{\varphi_0}e^{\frac{b_1}{2a_1}}e^{-\frac{s^2}{2a_1\varphi_0^2}}, s \to \infty, \end{aligned}$$

as was to be shown.

We now assume that the lemma is true for $k-1$. Since $q_k(s) = H(q_{k-1})(s)$, we have, by lemma 5.1, that

$$q_k(s) \simeq C' s^{(2\beta_{k-1}-\alpha_{k-1})/((\alpha_{k-1}+2))} e^{-c' s^{\frac{2\alpha_{k-1}}{\alpha_{k-1}+2}}}, \quad s \to \infty \tag{31}$$

Now $\frac{2\alpha_{k-1}}{\alpha_{k-1}+2} = 2/k = \alpha_k$ and, similary,

$$\frac{2\beta_{k-1} - \alpha_{k-1}}{\alpha_{k-1}+2} = -(1-1/k) = \beta_k,$$

using the expressions for β_{k-1} and α_{k-1}. From lemma 5.1 we get:

$$\begin{aligned} c' &= \frac{1}{2}(\alpha_{k-1}+2)c_{k-1}^{\frac{2}{\alpha_{k-1}+2}}(\alpha_{k-1}a_1)^{-\frac{\alpha_{k-1}}{\alpha_{k-1}+2}} \\ &= \frac{k}{2a_1\varphi_0^{2/k}} = c_k, \end{aligned}$$

after a short computation. Finally, by lemma 5.1 again,

$$\begin{aligned} C' &= \frac{2e^{\frac{b_1}{2a_1}}}{\sqrt{\alpha_{k-1}+2}}(c_{k-1}\alpha_{k-1}a_1)^{-\frac{\beta_{k-1}+1}{\alpha_{k-1}+2}} \\ &= 2e^{\frac{b_1}{2a_1}}\sqrt{\frac{k-1}{2k}}\left(\frac{1}{\varphi_0^{2/(k-1)}}\right)^{-1/2k}C_{k-1} \\ &= e^{\frac{b_1}{2a_1}}\sqrt{\frac{2(k-1)}{k}}\varphi_0^{1/k(k-1)}C_{k-1}, \end{aligned}$$

with

$$C_1 = \frac{1}{\sqrt{2\pi}}\frac{1}{a_1}\frac{1}{\varphi_0}e^{\frac{b_1}{2a_1}}$$

and a simple induction allows us to verify the formula for C_k. QED

We can now state the main results of this section:

Theorem 5.4 *Let $(r_t)_t$ be a GARCH(1, 1) process, with φ given by (27) and independent, normally distributed ε_t with mean 0 and variance 1. Fix a time t and a time-horizon $t+k$ and suppose that $r_t = \rho_0$ and $\sigma_t = s_0$ are given. Let $\varphi_0 := \varphi(\rho_0, s_0) = (a_0 + a_1\rho_0^2 + b_1 s_0^2)^{1/2}$, as before, and define constants c_k and C_k by*

$$c_k = \frac{1}{2}k a_1^{-(1-1/k)}\varphi_0^{-2/k},$$

and

$$C_k = \frac{e^{\frac{(k-1)b_1}{2a_1}}}{\sqrt{2\pi}}a_1^{-\frac{1}{2}(1-1/k)}\sqrt{\frac{2^{k-1}}{k}}\varphi_0^{-1/k}.$$

Then

$$p_{t,k}(x;\rho_0,s_0) \simeq C_k\frac{e^{-c_k|x|^{2/k}}}{|x|^{1-1/k}}, \quad x \to \pm\infty \tag{32}$$

Proof. The proof follows easily from the previous two lemmas. Using the notations introduced in the proof of lemma 5.3, we have that

$$p_{t,k}(x;\rho_0,s_0) = F(H^{k-1}(\delta_{\varphi_0})) = F(q_k(x)).$$

For $k=1$, we immediately get:

$$F(\delta_{\varphi_0})(x) = \frac{1}{\varphi_0\sqrt{2\pi}} e^{-\frac{x^2}{2\varphi_0^2}}.$$

For $k>1$ we have, by the previous lemma,

$$q_{k-1}(s) \simeq \tilde{C}_{k-1} s^{\beta_{k-1}} e^{-\tilde{c}_{k-1} s^{\alpha_{k-1}}}.$$

where $\tilde{C}_{k-1}$ and $\tilde{c}_{k-1}$ are the constants introduced in lemma 5.3. Therefore, by lemma 5.2,

$$p_k(x) = F(v)(x) \simeq C'|x|^{(2\beta_{k-1}-\alpha_{k-1})/(\alpha_{k-1}+2)} e^{-c' x^{\frac{2\alpha_{k-1}}{\alpha_{k-1}-2}}},$$

where C' and c' are the constants given by that lemma, with $c=\tilde{c}_{k-1}$ and $C=\tilde{C}_{k-1}$. After some computations we get that

$$c' = c_k = \frac{k}{2}\frac{1}{a_1^{(k-1)/k}}\frac{1}{\varphi_0^{2/k}},$$

and

$$C' = C_k = \frac{e^{\frac{(k-1)b_1}{2a_1}}}{\sqrt{2\pi}} a_1^{-\frac{1}{2}(1-1/k)} \sqrt{\frac{2^{k-1}}{k}} \varphi_0^{-1/k},$$

as stated. QED

Remarks 5.5 (i) It is surprising that the only difference between an ARCH(1) and a GARCH(1,1), as concerns the asymptotic behavior of their pdf, is the factor of $\exp((k-1)b_1/2a_1)$ in front.

(ii) The proof of theorem 5.4 will in fact give a complete asymptotic expansion, since lemma 4.1 does. A straightforward but somewhat tedious analysis, of which we will skip the details, shows that

$$p_{t,k}(x;\rho_0,s_0)\frac{e^{-c_k|x|^{2/k}}}{|x|^{1-1/k}} \sum_{\nu\geq 0} C_{\nu,k}|x|^{-\nu/k}, \quad |x|\to\infty$$

where $C_{0,k}=C_k$; a certain number of these coefficients (which can in principle all be computed explicitly, given sufficiently hard work) will vanish automatically.

6 Asymptotics of multi-period returns

If r_t represents a one-period logarithmic return, looking at r_{t+k} all by itself does not make much financial sense. A more relevant quantity would be the k-period return (17): $r_{t+k,t} = r_{t+1} + r_{t+2} + \cdots + r_{t+k}$. The main result of this section is that , qualitatively, the tails of the conditional pdf of $r_{t+k,t}$ behaves like that of r_{t+k} :

Theorem 6.1 *Let $(r_t)_t$ be a classical GARCH(1, 1), with $\varphi(r, \sigma) = (a_0 + a_1 r^2 + b_1 \sigma^2)^{1/2}$ and standard normally distributed $(\varepsilon_t)_t$, where we moreover suppose that $b_1 > 0$. Let $r_{t+k,t}$ be the k-period return defined above. Fix a k and let $\rho_0 \in \mathbf{R}, s_0 > 0$. Then there exist constants $c_k, c'_k, C_k, C'_k > 0$, depending on k, a_0, a_1, b_1, ρ_0 and s_0 such that for $|x| \geq 1$, say,*

$$C'_k |x|^{-(1-1/k)} e^{-c_k |x|^{2/k}} \leq \mathbf{P}(r_{t+k,t} = x | r_t = \rho_0, \sigma_t = s_0) \leq C_k |x|^{-(1-1/k)} e^{-c_k |x|^{2/k}} \tag{33}$$

Explicit values for the constants can be extracted from the proof below: we won't do that here. Note that we restricted ourselves to genuine GARCH(1, 1) processes. The proof, as it stands, would give a slightly worse result for an ARCH(1), in terms of the power of $|x|$ in front of the exponential in (33), but this is probably a technical problem.

Proof. The proof is based on formula (19) from section 3, which in our situation reads:

$$\mathbf{P}(r_{t+k,t} = x | r_t = \rho_0, \sigma_t = s_0) = \tag{34}$$
$$\left(\frac{1}{2\pi}\right)^{(k-1)/2} \int_{\mathbf{R}} \cdots \int_{\mathbf{R}} \Pi_{j=1}^{k-1} \frac{1}{\hat{s}_j} e^{-x_j^2/2\hat{s}_j^2} \frac{1}{\hat{s}_k} e^{-(x-(x_1+\cdots+x_{k-1}))^2/2\hat{s}_k^2} dx_1 \cdots dx_{k-1},$$

where the standard deviations $\hat{s}_j = \hat{s}_j(x_1, \cdots, x_{j-1})$ are defined inductively by

$$\hat{s}_1^2 = a_0 + a_1 \rho_0^2 + b_1 s_0^2$$
$$\hat{s}_j^2 = a_0 + a_1 x_{j-1}^2 + b_1 \hat{s}_{j-1}^2$$

It easily follows that

$$\hat{s}_j^2 = \sum_{\nu=1}^{j-1} a_1 b_1^{\nu-1} x_{j-\nu}^2 + e_\nu,$$

where $e_1 = \hat{s}_1^2$ and $e_k = a_0 + b_1 e_{k-1}$. We will in fact establish a slightly more general result, replacing the $\hat{s}_j^2$ in (34) by functions $L_{j-1} = L_{j-1}(x_1, \cdots, x_{j-1})$

which are affine in in $x_1^2, \cdots x_{j-1}^2$ (note the shift of the index by 1 in the notation here w.r.t. that used for $\hat{s}_j$). Here

$$L_j(x_1, \cdots, x_j) = \gamma_0^{(j)} + \sum_{\nu=1}^{j} \gamma_\nu^{(j)} x_\nu^2 \tag{35}$$

For our proof to work we will have to impose the condition:

$$\gamma_\nu^{(j)} > 0, \quad 0 \le \nu \le j \tag{36}$$

Note that L_0 is thus just a strictly positive constant. The $\hat{s}_j^2$ coming from a GARCH(1, 1) with $b_1 > 0$ fall into this class; those coming from an ARCH do not.

We will also put an adjustable multiplicative constant $\eta > 0$ in the exponent of the final factor of (34) and estimate the functions $q_k(x)$ defined by

$$q_k(x) = q_k(x; \eta, L_0, \cdots, L_{k-1}) = \tag{37}$$

$$\int_{\mathbf{R}} \cdots \int_{\mathbf{R}} \left(\Pi_{j=1}^{k-1} \frac{e^{-x_j^2/2L_{j-1}}}{\sqrt{2\pi L_{j-1}}} \right) \frac{1}{\sqrt{2\pi L_{k-1}}} e^{-\eta(x-(x_1+\cdots x_{k-1}))^2/2L_{k-1}} dx_1 \cdots dx_{k-1}.$$

More precisely, we will prove the following inequalities, from which theorem 6.1 will be an immediate consequence:

Claim 6.2 *For given k, affine forms $L_0, \cdots L_{k-1}$ as in (35), satisfying (36) and given $\eta > 0$, there exist strictly positive constants c, c', C and C' such that*

$$C|x|^{-(1-1/k)} e^{-c|x|^{2/k}} \le q_k(x) \le C'|x|^{-(1-1/k)} e^{-c'|x|^{2/k}} \tag{38}$$

The constants c, c', C and C' can be chosen locally uniformly in η and in (the coefficients of) the L_j.

We turn to the proof of the claim, which will be by induction on k. The idea is to estimate $q_k(x)$ from above and from below by a Laplace transform of a q_{k-1} with slightly modified η and L 's, modulo a negligible error, and then use lemma 4.1 again. To accomplish this, we will eliminate the x_1 from all factors under the integral sign of (37), except the first one. For this we use the following elementary inequality:

Lemma 6.3 *For all ε with $0 < \varepsilon \leq 1$ and all $a, b \in \mathbf{R}$ one has that*

$$C^-_{b,\varepsilon} e^{-(1+\varepsilon)a^2} \leq e^{-(a+b)^2} \leq C^+_{b,\varepsilon} e^{-(1-\varepsilon)a^2}, \tag{39}$$

where $C^-_{b,\varepsilon} = \exp(-(\varepsilon^{-1}+1)b^2)$ and $C^+_{b,\varepsilon} = \exp((\varepsilon^{-1}-1)b^2)$

Proof. To prove for example the upper bound, write $\exp((1-\varepsilon)a^2)\exp(-(a+b)^2) = \exp -(\varepsilon a^2 + 2ab + b^2)$ and maximize over a. The lower bound is proven in the same way.

It is clear that (38) holds for $k = 1$. Now suppose that it holds for $k-1$. We then have to prove it for k. We first establish the upper bound in (38). Apply the second inequality in (39) with $a = \sqrt{\eta}(x - (x_2 + \cdots + x_k))/\sqrt{2L_{k-1}}$ and $b = -\sqrt{\eta}x_1/\sqrt{2L_{k-1}}$. The constant $C^+_{b,\varepsilon}$ then becomes

$$C^+_{b,\varepsilon} = e^{(\varepsilon^{-1}-1)\eta x_1^2/2L_{k-1}} \leq e^{(\varepsilon^{-1}-1)\eta x_1^2/2\gamma_0^{(k-1)}},$$

and we see that it can be absorbed in the first factor in the integrant of (37), $\exp(-x_1^2/2L_0)$, provided ε is sufficiently close to 1. In fact, $C_{b,\varepsilon} < \exp x_1^2/4L_0$ if

$$(1 + \gamma_0^{(k-1)}/2\eta L_0)^{-1} < \varepsilon < 1.$$

With this choice of ε we then have that

$$q_k(x) \leq \int_{\mathbf{R}} \cdots \int_{\mathbf{R}} \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \cdot \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2L_{j-1}}}{\sqrt{2\pi L_{j-1}}}\right) \cdot \tag{40}$$

$$\cdot \frac{1}{\sqrt{2\pi L_{k-1}}} e^{-(1-\varepsilon)\eta(x-(x_2+\cdots x_{k-1}))^2/2L_{k-1}} dx_1 \cdots dx_{k-1}.$$

We now split this integral as

$$\int_{|x_1|\leq 1} dx_1 \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \int_{\mathbf{R}^{k-2}} (\cdots) + \int_{|x_1|>1} dx_1 \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \int_{\mathbf{R}^{k-2}} (\cdots) \tag{41}$$

$$= I + II$$

and estimate the two pieces separately. We first show that I is of the same order as a $q_{k-1}(x; \eta^*, L_1^*, \cdots, L_{k-1}^*)$ for some suitable choice of η^* and L_ν^*. In fact, if $|x_1| \leq 1$, then for $\nu \geq 1$,

$$\begin{aligned} L_j(x_1, \cdots x_j) &\leq (\gamma_0^{(j)} + \gamma_1^{(j)} + \gamma_2^{(j)} x_2^2 + \cdots + \gamma_j^{(j)} x_j^2) \\ &=: L_j^*(x_2, \cdots, x_j), \end{aligned}$$

where L_1^* will just be a constant, independent of $x_2, \cdots, x_{k-1}$. We also have that

$$\frac{L_j^*}{L_j} \leq \max(1, (\gamma_0^{(j)} + \gamma_1^{(j)})/\gamma_0^{(j)}),$$

this without any restriction on $(x_1, \cdots, x_j)$. It follows that, for a suitable constant $C > 0$,

$$|I| \leq C \int_{|x_1| \leq 1} \int_R \cdots \int_R \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \cdot \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2L_{j-1}^*}}{\sqrt{2\pi L_{j-1}^*}} \right)$$

$$\cdot \frac{1}{\sqrt{2\pi L_{k-1}^*}} e^{-(1-\varepsilon)\eta(x-(x_2+\cdots x_{k-1}))^2/2L_{k-1}^*} dx_1 \cdots x_{k-1}.$$

We recognize the integral over $dx_2 \cdots dx_{k-1}$ as a constant times $q_{k-1}(x;(1-\varepsilon)\eta; L_2^*, \cdots, L_{k-1}^*)$, and therefore, by the induction hypothesis, for suitable constants c, C,

$$|I| \leq C|x|^{-(1-1/(k-1))} e^{-c|x|^{2/(k-1)}}, \tag{42}$$

which for $|x| \to \infty$ is of strictly lower order than the inequality we're trying to establish for $q_k(x)$.

We next turn to the integral II. If $|x_1| > 1$, then

$$\begin{aligned} L_j(x_1, \cdots, x_j) &\leq (\gamma_0^{(j)} + \gamma_1^{(j)}) x_1^2 + \gamma_2^{(j)} x_2^2 + \cdots + \gamma_j^{(j)} x_j^2 \\ &= x_1^2 \left(\gamma_0^{(j)} + \gamma_1^{(j)} + \gamma_2^{(j)} \frac{x_2^2}{x_1^2} + \cdots + \gamma_j^{(j)} \frac{x_j^2}{x_1^2} \right) \\ &=: x_1^2 \widetilde{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right), \end{aligned}$$

the last equation defining $\widetilde{L}_j$. Similarly, for $|x_1| > 1$ we can estimate

$$\begin{aligned} L_j(x_1, \cdots, x_j) &\geq \gamma_1^{(j)} x_1^2 + \gamma_2^{(j)} x_2^2 + \cdots + \gamma_j^{(j)} x_j^2 \\ &= x_1^2 \left(\gamma_1^{(j)} + \gamma_2^{(j)} \frac{x_2^2}{x_1^2} + \cdots + \gamma_j^{(j)} \frac{x_j^2}{x_1^2} \right) \\ &\geq c x_1^2 \widetilde{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right), \end{aligned} \tag{43}$$

$$\tag{44}$$

provided that

$$c \leq \frac{\gamma_1^{(j)}}{\gamma_0^{(j)} + \gamma_1^{(j)}}.$$

Note that to have (43) with a $c > 0$ we need here that $\gamma_1^{(j)} > 0$, which is assured by condition (36). Substituting these inequalities in (37), we find that for suitable $C > 0$,

$$II \leq C \int_{|x_1|>1} \int_{\mathbf{R}^{k-2}} \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2x_1^2 \widetilde{L}_{j-1}}}{|x_1| \sqrt{2\pi \widetilde{L}_{j-1}}} \right) \cdot$$

$$\cdot \frac{1}{|x_1| \sqrt{2\pi \widetilde{L}_{k-1}}} e^{-(1-\varepsilon)\eta(x-(x_2+\cdots x_{k-1}))^2/2x_1^2 \widetilde{L}_{k-1}} dx_1 \cdots dx_{k-1}$$

(we can in fact take $C = \left(\min_j \gamma_1^{(j)} / (\gamma_0^{(j)} + \gamma_1^{(j)}) \right)^{-(k-1)/2}$). If we now change variables to $y_j := x_j / |x_1|$, $2 \leq j \leq k-1$ we see that the previous inequality can be written as:

$$II \leq C \int_{|x_1|>1} \frac{1}{|x_1|} \frac{e^{-x_1^2/4L_0}}{\sqrt{2\pi L_0}} q_{k-1} \left(\frac{x}{|x_1|};\ (1-\varepsilon)\eta, \widetilde{L}_1, \cdots, \widetilde{L}_{k-1} \right) dx_1.$$

By the induction-hypothesis, the $q_{k-1}(x/x_1)$ under the integrant is less than or equal to

$$C \left(\frac{|x|}{|x_1|} \right)^{\beta} e^{-c(|x|/|x_1|)^{\alpha}}$$

with

$$\alpha = 2/(k-1), \quad \beta = -1 + 1/(k-1) \tag{45}$$

and thus, after a rescaling, and with different constant c and C,

$$II \leq |x|^{\beta} C \int_{|x_1|>1} |x_1|^{-\beta-1} e^{-c(|x|/|x_1|)^{\alpha}} e^{-x_1^2} dx_1.$$

We now assume that $x > 0$ and we write the integral as twice the integral over $[1, \infty)$. We again want to use lemma 4.1 and for this we rewrite our integral as a Laplace transform with big parameter, by introducing the new variable $z = x_1^{-\alpha}$. Then the right hand side of (46) is less or equal a constant times

$$x^{\beta} \int_0^1 z^{(\beta/\alpha)-1} e^{-z^{-2/\alpha}} e^{-cx^{\alpha} z} dz$$

and by lemma 4.1, with $s = cx^{\alpha}$ and α, β replaced by, respectively, $2/\alpha$ and $1 - (\beta/\alpha)$, we find that (again with a different c and C),

$$II \leq C x^{(2\beta-\alpha)/(\alpha+2)} \exp(-c x^{2\alpha/(\alpha+2)}).$$

Since, using (45), the exponents of x in this formula turn out to be, respectively, $-(1-1/k)$ and $2/k$, this proves the desired upper bound for II and thus for $q_k(x)$.

We next turn to the lower bound for q_k. By the first inequality of lemma (39), we see in the same way as before that

$$e^{-\eta(x-(x_1+\cdots+x_{k-1}))^2/2L_{k-1}} \geq C^-_{b,\varepsilon} e^{-(1+\varepsilon)\eta(x-(x_2+\cdots+x_{k-1}))^2/2L_{k-1}}$$

where

$$C^-_{b,\varepsilon} = e^{-\eta(1+\varepsilon^{-1})x_1^2/2L_{k-1}} \geq e^{-\eta(1+\varepsilon^{-1})x_1^2/2\gamma_0^{(k-1)}}.$$

We can combine $C_{b,\varepsilon}$ with the first factor of the integrand of the defining equation (37) of $q_k(x)$ into a factor $e^{-\kappa x_1^2}$. Doing so, and limiting the x_1-integration in (37) to $|x_1| > 1$, we find that

$$q_k(x) \geq \int_{|x_1|>1} \int_R \cdots \int_R \frac{e^{-\kappa x_1^2}}{\sqrt{2\pi L_0}} \cdot \left(\Pi_{j=2}^{k-1} \frac{e^{-x_j^2/2L_{j-1}}}{\sqrt{2\pi L_{j-1}}} \right) \tag{46}$$

$$\cdot \frac{1}{\sqrt{2\pi L_{k-1}}} e^{-\eta(1+\varepsilon)(x-(x_2+\cdots x_{k-1}))^2/2L_{k-1}} dx_1 \cdots x_{k-1}.$$

As before, we next get rid of the x_1 in the $L_1, \cdots, L_{k-1}$; first, if $j \geq 1$, then

$$\begin{aligned} L_j(x_1, \cdots x_j) &\geq x_1^2 \left(\gamma_1^{(j)} + \gamma_2^{(j)} \frac{x_2^2}{x_1^2} + \cdots + \gamma_j^{(j)} \frac{x_j^2}{x_1^2} \right) \\ &=: x_1^2 \hat{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right). \end{aligned}$$

Next, if $|x_1| > 1$, then

$$\begin{aligned} L_j(x_1, \cdots, x_j) &\leq (\gamma_0^{(j)} + \gamma_1^{(j)}) x_1^2 + \cdots + \gamma_j^{(j)} x_j^2 \\ &\leq c x_1^2 \hat{L}_j \left(\frac{x_2}{x_1}, \cdots, \frac{x_j}{x_1} \right), \end{aligned} \tag{47}$$

provided that $c \geq (\gamma_0^{(j)} + \gamma_1^{(j)})/\gamma_1^{(j)}$; there exists such a (finite) c since $\gamma_1^{(j)} > 0$ by (36). Substituting these inequalities in (46) and making the same change of variables $y_j = x_j/x_1$ as before ($j \geq 2$) one finds that, for a suitable constant $C > 0$,

$$q_k(x) \geq C \int_{|x_1|>1} \frac{e^{-\kappa x_1^2}}{|x_1|} q_{k-1} \left(\frac{x}{x_1}; (1+\varepsilon)\eta, \hat{L}_2, \cdots, \hat{L}_{k-1}, \eta(1+\varepsilon) \right).$$

Using the induction hypothesis and lemma 4.1, we find the required lower bound for $q_k(x)$. QED

7 Application to quantile estimation and VaR

The asymptotic estimates from the previous two sections are, in principle, relevant for the analysis and forecasting of financial risk. We will give an illustration of this in the present section, by applying them to the estimation of multi-period Value at Risk in GARCH(1, 1) models, and, in particular, by comparing the latter with the well-known heuristical "$\sqrt{\text{period}}$-rule" mentioned in the introduction.

From a mathematical point of view, VaR is of course just a lower quantile of the probability distribution for the Profit & Loss function over the period in which one is interested, where it is a matter of some discussion whether this probability should be taken unconditionally or conditionally to, e.g., price information available up to time t (cf. for example [15]). We will take here the conditional point of view. Recall that if $F_X = \mathbf{P}(X \leq \cdot)$ is the cumulative distribution function of some random variable X, then its quantile function $F_X^{\leftarrow}$ can be defined as:

$$F^{\leftarrow}(\alpha) := \inf\{x : F(x) = \alpha\}, \quad \alpha \in [0, 1]; \tag{48}$$

cf. for example Embrechts *et al.* [12], definitions 3.3.4 and 3.3.5. We observe that if F_X is continuous, than $F(F^{\leftarrow}(\alpha)) = \alpha$, but not necessarily so at points of discontinuity. Thus for continuous F_X, $\mathbf{P}(X \leq F^{\leftarrow}(\alpha)) = \alpha$, or, equivalently: with probability $1 - \alpha$, $X > F^{\leftarrow}(\alpha)$. We recognize the financial interpretation of VaR.

If the probability $\mathbf{P}(\cdot)$ is taken conditionally with respect to (some value y of) some other (vector-) random variable Y, we denote the conditional cdf of X by $F_{X|Y}$ ($F_{X|Y=y}$). Now Let $(P_t)_{t\in\mathbf{N}}$ be some discrete parameter stochastic process of non-negative random variables, to be interpreted as the time series of prices of some financial asset. The *(conditional) Value at Risk at time t, with confidence $1 - \alpha$ over the time window* $[t, t+k]$ of the asset then is simply defined as:

$$VaR_{1-\alpha}(t, t+k) := F^{\leftarrow}_{P_{t+k}-P_t|P_t}(\alpha). \tag{49}$$

Note that this is equivalent to the informal definition in words given in the introduction, at least if the distribution function $F_{P_{t+k}|P_t}$ is continuous, as it will be in our case. Also note that a loss is recorded here as a negative number. To fix ideas, we have conditioned here on the price at t, but one might add other random variables, as we will do below by in the context of GARCH-processes by including the volatility.

Usually prices are modeled through the process of daily log-returns: $r_t = \log(P_t/P_{t-1})$, and we consequently condition on r_t instead of P_t. A short computation shows that

$$VaR_{1-\alpha}(t, t+k) = (\exp(F^{\leftarrow}_{r_{t+k,t}|r_t}(\alpha)) - 1)P_t. \tag{50}$$

Note that the maximum loss over $[t, t+k]$ is necessarily bounded from below by (minus) once's fortune at t. If $F^{\leftarrow}_{r_{t+k,t}|r_t}$ is small, as in practice it usually is (or should be!), one simply approximates this by $F^{\leftarrow}_{r_{t+k,t}|r_t,\sigma_t}(\alpha))P_t$.

We will suppose from now on that $(r_t)_t$ is given by a (classical) GARCH(1, 1), as in the previous two sections. As announced, we now also condition on the volatility σ_t at time t, which is known in this model. Recall that the complementary error function Erfc can be defined as:

$$\text{Erfc }(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2}.$$

Theorem 6.1 then easily implies, upon integration and using symmetry, that for $x < 0$,

$$F_{r_{t+k,t}|r_t,\sigma_t}(x) \leq k\pi^{1/2} c_k^{-1/2} C_k \cdot \text{Erfc }(\sqrt{2c_k}|x|^{1/k}), \tag{51}$$

and a similar lower bound valid for $x < -1$, say, with the constants replaced by the primed ones. We now use these estimates, together with the well-known asymptotics of the complementary error function,

$$\text{Erfc }(x) \simeq \frac{1}{\sqrt{2\pi}} \frac{e^{-x^2/2}}{x}, \quad x \to \infty,$$

to give asymptotic bounds $F^{\leftarrow}_{r_{t+k,t}|r_t,\sigma_t}(\alpha)$. We will proceed informally and simply replace Erfc $(\cdot)$ by it's asymptotic expression above; however, all of the following can done quite rigorously. Making this approximation, the sought-for quantile can then be bounded, asymptotically for $\alpha \to 0$, by:

$$\rho'(\alpha) \leq F^{\leftarrow}_{r_{t+k,t}|r_t=\rho_0,\sigma_t=s_0}(\alpha) \leq \rho(\alpha), \tag{52}$$

where $\rho(\alpha)$ is the solution of

$$(\text{Constant}) \frac{\exp(-c_k \rho(\alpha)^{2/k})}{(2c_k \rho(\alpha)^{2/k})^{1/2}} = \alpha,$$

with some constant in front which we won't need to know; $\rho'(\alpha)$ will be the solution of a similar equation, with different constants, and in particular with c_k replaced by c'_k. Taking logarithms, we see that we have to study the transcendental equation

$$z + \frac{1}{2} \log z = w, \tag{53}$$

with

$$z = c_k \rho(\alpha)^{2/k},$$

and

$$w = \log \alpha^{-1} + \text{Constant}',$$

with some different constant. Now if α is sufficiently close to 0, the solution $z = z(w)$ will be bigger than 1, and therefore $\log z > 0$. It then follows that $z < w$, and therefore $\log z < \log w$. Hence $z = w - (\log z)/2 > w - (\log w)/2$. This showes that

$$w - (\log w)/2 < z < w. \tag{54}$$

(This elementary argument, which we learned from Pierre Duclos, can be iterated to obtain an expansion of z in terms of w.) It easily follows from this that

$$\frac{\rho(\alpha)}{(\log \alpha^{-1})^{k/2}} \to c_k^{-k/2}, \quad \alpha \to 0.$$

If we repeat the same argument for $\rho'(\alpha)$, we arrive at the following result:

Proposition 7.1 *Suppose that $(r_t)_t$ is a classical GARCH(1, 1) as in example 2.3(i), with normal innovations, and let $F_k^{\leftarrow} = F^{\leftarrow}_{r_{t+k}|r_t=\rho_0,\sigma_t=s_0}$ be the conditional quantile function. Then we have that for each fixed $k \geq 1$,*

$$(c')_k^{-k/2} < \liminf_{\alpha \to 0} \frac{F_k^{\leftarrow}(\alpha)}{(\log \alpha^{-1})^{k/2}} \leq \limsup_{\alpha \to 0} \frac{F_k^{\leftarrow}(\alpha)}{(\log \alpha^{-1})^{k/2}} < c_k^{-k/2},$$

where c_k, c_k' are the constants from theorem 6.1.

Note that in particular for $k = 1$ we have that $F_k^{\leftarrow}(\alpha)$ behaves roughly like $(\log \alpha^{-1})^{1/2}$ (in this case one can of course easily prove much more precise estimates, since the conditional distribution of r_{t+1} is just a normal one). Hence we have the following

Corollary 7.2 *(comparaison with the $\sqrt{k}$-rule). For each $k \geq 2$,*

$$\lim_{\alpha \to 0} \frac{F_k^{\leftarrow}(\alpha)}{\sqrt{k} F_1(\alpha)} = \infty.$$

So we see that for small α the $\sqrt{k}$-rule may fail spectacularly, even for a k as small as 2.

It is natural to ask for the best values of the constants in proposition 7.1. For the constant c_k in the upper bound we can easily find an explicit value from theorem 5.4. To see this, first note that if $\lambda_1 + \cdots + \lambda_k = 1$, $\lambda_j \geq 0 (1 \leq j \leq k)$, then for any $x \in \mathbf{R}$,

$$\{r_{t+k,t} < x\} \subset \cup_{j=1}^{k} \{r_{t+j} < \lambda_j x\},$$

and therefore

$$\mathbf{P}(r_{t+k,t} < x) \leq \sum_{j=1}^{k} \mathbf{P}(r_{t+j} < \lambda_j x). \tag{55}$$

Now, by the same computation as the one leading to (53), we find using theorem 5.4 that

$$\mathbf{P}(r_{t+j} < x) \simeq \widetilde{C}_j \mathrm{Erfc} \left(\sqrt{j} a_1^{-(1-1/j)} \varphi_0^{-1/j} |x|^{1/j} \right), \tag{56}$$

where the constant can be evaluated as $\widetilde{C}_j = \exp((j-1)(\log 2 + b_1/a_1)/2)$ (although we won't use the exact value here, except to note that it is non-zero). Inserting this into the right hand side of (55) and taking the infimum over all allowed choices for $\lambda_1, \cdots, \lambda_k$ will give an upper bound for $F_{r_{t+k,t}|r_t,\sigma_t}$. Here we will simply choses $\lambda_k = 1 - \varepsilon$, $\lambda_j = \varepsilon/(k-1)$ for $1 \leq j \leq k-1$, for some $0 < \varepsilon < 1$ which will tend to 0 at the end. A moment's thought then shows that for all $\eta > 0$ there exists an $R = R(\eta, \varepsilon, k) > 0$ such that for all $x < -R$,

$$F_{r_{t+k,t}|r_t=\rho_0,\sigma_t=s_0}(x) \leq (1+\eta) \widetilde{C}_k \mathrm{Erfc} \left((1-\varepsilon) \sqrt{k} a_1^{-(1-1/k)} \varphi_0^{-1/k} |x|^{1/k} \right).$$

(We note in passing that this gives another proof of (51), one which does not depend on theorem 6.1; however, the latter is a result on the probability density, which cannot be derived from estimates on the cumulative distribution function.)

Using this inequality as before to bound the α-th quantile of the left hand side from above, one finds that

$$\limsup_{\alpha \to 0} \frac{F^{\leftarrow}_{r_{t+k,t}|r_t=\rho_0,\sigma_t=s_0}(\alpha)}{(\log \alpha^{-2})^{k/2}} \leq (1-\varepsilon)^{-k} k^{-k/2} a_1^{-(k-1)} \varphi_0. \tag{57}$$

Letting $\varepsilon \to 0$, we conclude that:

Proposition 7.3 *Using the same notations as in proposition 7.1 we have that*

$$\limsup_{\alpha \to 0} \frac{F_k^{\leftarrow}(\alpha)}{(\log \alpha^{-2})^{k/2}} \leq k^{-k/2} a_1^{-(k-1)} \varphi_0, \tag{58}$$

where we recall that $\varphi_0 = (a_0 + a_1 \rho_0^2 + b_1 s_0^2)^{1/2} = \sigma_{t+1}$, *the volatility over* $[t, t+1]$ *in our GARCH(1, 1) model.*

We conjecture that the limsup in (58) is actually a limit, and equal to the right hand side of this inequality.

Note that, somewhat counter-intuitively, the constant on the right of inequality (58) decreases quite rapidly for increasing k. This should be taken as an indication that one should be carefull with using (58) to draw conclusions about numerical values of $VaR_\alpha(t+k,t)$ for a given α and k: (58) is an asymptotic inequality which will only become valid for sufficiently small values of α, where the precise meaning of "sufficiently small" will depend on k. We believe that the asymptotic results on GARCH(1, 1) models and their implications for VaR are in the first place of theoretical interest. For the ranges of $\alpha = 00.1 - 0.05$ and $k = 1 - 10$ which are relevant for the practice of risk management, one will probably be better advised using numerical implementation of the semi-explicit formulas from sections 2 and 3. In particular, for finding a quick upper bound for one's VaR, one could use the numerically very efficient formula (9) in conjunction with (55) above; finding an optimal (in general x-dependent) choice of the λ_j will of course be an issue here. We refer to [6] for a further discussion of these and other numerical issues.

References

[1] Artzner, Ph., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent measures of risk, Mathematical Finance, 9 (3), 203 - 228.

[2] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity, J. of Econometrics 31, 307 - 325.

[3] Bollerslev, T., Engle, R. F., Nelson, D. B., 1994. ARCH models, in: Engle, R. F., McFadden, D., eds., *Handbook of Econometrics, vol. 4*, North Holland, Amsterdam.

[4] Bingham, N. H., Goldie, G. M., Teugels, J. L.,1987. *Regular Variation*, Cambridge University Press, Cambridge.

[5] Brummelhuis, R. G. M., Guégan, D., 2000. Extreme values of conditional distributions of GARCH(1, 1) processes, prépublication 00.08, Université de Reims, Dépt. de math., submitted.

[6] Brummelhuis, R. G. M. , Guégan, D. and LaDoucette, S. In preparation.

[7] Diebolt, J., Guégan, D. , 1991, Le modèle β-ARCH, Comptes Rendues Acad. Sciences Paris, Série I, 312, 625 - 630.

[8] Diebolt, J., Guégan, D. , 1993, Tail behavior of the stationary density of general non linear autoregressive process of order one, J. Appl. Prob., 30, 315 - 329.

[9] Dingh, Z., Granger, C.W.J., 1996, Modelling volatility persistence of speculative returns: a new approach, Journal of Econometrics, 73, 185 - 215.

[10] Drost, F. C., Nijman, T. E.,1993, Temporal aggregation of GARCH processes, Econometrica, 61, 909 - 927.

[11] Dowd, K.,1998, *Beyond Value at Risk*, Wiley Frontiers in Finance.

[12] Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[13] Engle, R. F., 1982. Autoregressive conditional heteroskedacity with estimates of the variance of United Kingdom inflation, Econometrica 50, 907 - 1007.

[14] Engle, R. F., Lilian, D. F., Robbins, R. P., 1987. Estimating time varying risk premia in the term structure, Econometrica 55, 391 - 407.

[15] Frey, R., McNeil, A., 2000, Estimation of tail-related risk measures for heteroscedastic financial time series : an extreme value approach, J. Empirical Finance, 7, 271 - 300.

[16] Guégan, D., 1994. *Séries chronologiques non linéaires à temps discrèt*, Economica, coll. Statistique mathématique et probabilité, Paris.

[17] Guégan, D., 2000. A new model: the k-factor GIGARCH process, Journal of Signal Processing 4, 265 - 271.

[18] Jorion, P., 1997. Value at Risk : The New Benchmark for Controlling Market Risk, McGraw-Hill.

[19] Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices, Acta Math. 131, 207 - 248.

[20] Nelson, D. B., 1990. Conditional heteroscedasticity in asset returns: a new approach, Econometrica 59, 347 - 370.

[21] Mikosch, Th., Starica, C., 2000. Limit theory for the sample autocorrelations and extremes of a GARCH(1, 1) process, Ann. Statistics 28 (5), 1427 - 1451.

[22] RiskMetrics - Technical Document, 1996, 4th edition, J. P. Morgan.

[23] Sloan, I. H., Joe, S.,*Lattice Methods for Multiple Integration*, 1994. Oxford Science Publications, Clarendon Press, Oxford.